

APPELLANT'S BRIEF UNDER 37 CFR §41.37		Docket No. C0012/7000
Applicant:	Raju C. Bopardikar, Jack J. Stiffler, Jacob Y. Bast, Gary A. Cardone, David E. Kaufman, Stuart P. MacEachern, Bruce D. McLeod, James M. Nolan, Jr., Zdenek Radouch and James A. Wentworth, III	
Serial No:	09/608,521	
Filed:	June 30, 2000	
For:	METHOD AND APPARATUS FOR IMPLEMENTING HIGH-PERFORMANCE, SCALEABLE DATA PROCESSING AND STORAGE SYSTEMS	
Examiner:	A. M. Mirza	
Art Unit:	2145	

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

This brief is in furtherance of the Notice of Appeal, filed in this case on April 30, 2009.

The fees required under §41.20(b)(2), and any required petition for extension of time for filing this brief and fees therefor, are dealt with in the accompanying TRANSMITTAL OF APPEAL BRIEF.

A single copy of this brief is transmitted (37 C.F.R. §41.37(a)) and contains these items under the following headings, and in the order set forth below (37 C.F.R. §41.37(c)(1)):

- I REAL PARTY IN INTEREST
- II RELATED APPEALS AND INTERFERENCES
- III STATUS OF CLAIMS
- IV STATUS OF AMENDMENTS
- V SUMMARY OF CLAIMED SUBJECT MATTER
- VI GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL
- VII ARGUMENT
- VIII APPENDIX OF CLAIMS INVOLVED IN THE APPEAL
- IX EVIDENCE APPENDIX
- X RELATED PROCEEDINGS APPENDIX

I REAL PARTY IN INTEREST (37 C.F.R. §41.37(c)(1)(i))

The real party in interest in this appeal is EMC Corporation, 176 South Street, Hopkinton, MA 01748.

II RELATED APPEALS AND INTERFERENCES (37 C.F.R. §41.37(c)(1)(ii))

There are no other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in the pending appeal.

III STATUS OF CLAIMS (37 C.F.R. §41.37(c)(1)(iii))

A. TOTAL NUMBER OF CLAIMS IN APPLICATION

The claims in the application are: 1-86.

B. STATUS OF ALL THE CLAIMS IN APPLICATION

1. Claims pending: 1, 3-11, 44 and 46-54.
2. Claims canceled: 2 and 45.
3. Claims withdrawn from consideration, but not canceled: 12-43 and 55-86.
4. Claims allowed: None.
5. Claims rejected: 1, 3-11, 44 and 46-54.

C. CLAIMS ON APPEAL

The claims on appeal are: 1, 3-11, 44 and 46-54.

IV STATUS OF AMENDMENTS (37 C.F.R. §41.37(c)(1)(iv))

No amendments have been filed.

V SUMMARY OF CLAIMED SUBJECT MATTER (37 C.F.R. §41.37(c)(1)(v))

The present invention relates to a method and apparatus for providing storage services to clients from a "pool" of storage devices (see page 4, line 21 - page 5, line 13 and Figure 1). The clients (102, 104) are connected to access interfaces (106, 108, 122) which are, in turn, connected to the pool of storage devices (118, 120, 122, 124) by a switch fabric (114, 116). The access interfaces communicate with the devices in the storage device pool and, based on the workload of each device in the storage device pool, select a subset of the storage devices to use for any given transaction and distribute the workload (page 5, lines 16-22). Thus, the data received from a client is transferred to the selected subset of storage devices based on the workload of the devices rather than on a physical device address (page 5, line 23 – page 6, line 6). The

result of this operation is that data may be stored in a storage device other than the storage device from which that data was retrieved (page 7, line 22 – page 8, line 11).

The correspondence between data identification and the location at which the data is stored is broken by using data “handles”. In particular, in the present system, a client wishing to retrieve or store data specifies a data object (a set of data) and a data identifier or “handle” for that object instead of a traditional physical address that is associated with a location (page 10, lines 11-25). In the case of a read operation, this handle is used to locate a map (called an “fmap”) which specifies the actual physical locations of the data (page 10, lines 16-25). In the case of a write operation, the data is stored in physical addresses specified by allocation units which are preallocated to each interface module by the storage devices before any write operation takes place (page 12, lines 11-18). The allocation units which are used for a particular data storage operation are selected based on the recent activity of the storage device with which they are associated (page 14, lines 12-18). After the write operation is committed, the fmaps associated with the data handle are updated to reflect the new location of the data (page 13, lines 11-14).

Therefore, during a typical read-modify-write operation, data corresponding a data identifier would be read from a storage location and modified, but might be stored back into a location different from which it was read depending on resource load and availability. The data thus dynamically moves based on criteria determined by the data storage system. Consequently, load balancing becomes automatic (page 14, line 19- page 15, line 2).

Therefore, data is not statically assigned to specific storage systems. Instead, when a storage device temporarily becomes overloaded by multiple clients simultaneously attempting to access the device, the client data can be automatically directed to, and stored in, another device. Further, the system can be easily scaled by simply adding new storage devices. Since these new storage devices have no usage when they are installed, data from overloaded devices will automatically be redirected to, and stored in, these devices. No reconfiguration of the system is required.

The independent claims correspond to the specification as follows:

Claim 1 - an access interface module (page 5, lines 9-14)

which receives from the client data storage requests, each including a data object to be stored and a data object identifier that identifies that data object (page 10, lines 11-25) and,

in response to each storage request and based on a workload and on relative demands placed on subsets of the plurality of storage devices instead of a physical location in the plurality of devices, dynamically selects a subset of the plurality of storage devices in which the data is stored (page 5, lines 16-22 and page 14, lines 12-18, page 18, lines 12-24)

so that data corresponding to the same data object identifier can be transferred to different physical storage device locations from request to request in order to dynamically distribute the workload across the plurality of storage devices (page 14, line 19-page 15, line 2); and

a switch fabric for temporarily connecting the access interface module to the selected subset of the plurality of storage devices so that the data can be transferred to the selected subset of storage devices (page 4, line 25 – page 5, line 8).

Claim 44

step (a) - page 10, lines 11-25

step (b) - page 5, lines 16-22 and page 14, lines 12-18; page 14, line 19-
page 15, line 2

step (c) - page 4, line 25 – page 5, line 8

The dependent claims correspond to the specification as follows:

Claim 3 – page 4, line 25 – page 5, line 1, Figure 9, 952 and 942.

Claim 4 - page 5, lines 1-8.

Claim 5 – page 10, lines 4-10 Figure 3, 302, 304.

Claim 6 – page 22, lines 21-28, Figure 9, 952 and 942.

Claim 7 – page 12, lines 11-15 Figure 2 218, 220, 222, 224).

Claim 8 – page 12, lines 15-17 , Figure 2, 206, 216, 218.

Claim 9 – page 12, lines 17-18.

Claim 10 – page 9, lines 14-18, Figure 3, 322.

Claim 11 – page 4, lines 23-24, Figure 1 106, 108, 110, 112.

Claim 46 – page 4, line 25 – page 5, line 1, Figure 9, 952 and 942.

Claim 47 - page 5, lines 1-8.

Claim 48 – page 10, lines 4-10 Figure 3, 302, 304.

Claim 49 – page 22, lines 21-28, Figure 9, 952 and 942.

Claim 50 – page 12, lines 11-15 Figure 2 218, 220, 222, 224).

Claim 51 – page 12, lines 15-17 , Figure 2, 206, 216, 218.

Claim 52 – page 12, lines 17-18.

Claim 53 – page 9, lines 14-18, Figure 3, 322.

Claim 54 – page 4, lines 23-24, Figure 1 106, 108, 110, 112.

VI GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL (37 C.F.R. 41.37 (c)(1)(vi))

Whether claims 1, 3-11, 44 and 46-54 are patentable under 35 U.S.C. 103(a) over U.S. Patent No. 6,195,703 (Blumenau) in view of U.S. Patent No. 5,909,686 (Muller.)

VII ARGUMENT (37 C.F.R. §41.37(c)(1)(vii))

Prima facie obviousness has not been established because the combination of Blumenau and Muller does not teach or suggest the structure recited in claims 1, 3-11, 44 and 46-54.

Obviousness is a legal conclusion based on factual evidence. Graham v. John Deere Co. 383 US 1, 148 USPQ 459 (1966). To establish the *prima facie* obviousness of a claimed invention, all of the claim limitations must be taught or suggested by the prior art. In re Wada and Murphy, Appeal 2007-3733, citing In re Ochiai, 71 F.3d 1565, 1572 (Fed. Cir. 1995) and In re Royka, 490 F.2d 981, 180 USPQ 580 (CCPA 1970).

The examiner refers to column 2, lines 5-12, column 2, lines 66-67 and column 3, lines 1-22 of the Blumenau reference. There the reference discloses a storage system in which a host can send a request to access storage resources through a switch to one of a plurality of network ports of a data storage subsystem. The switch allows paths to be selected between a particular host and alternative storage subsystem ports based on load balancing considerations. However, once a request arrives at any network port of the storage subsystem a data address in the request is translated to a physical location on one of the storage devices and the data passes through the port and is written to, or read from, that storage device. For example, Blumenau, column 4, lines 43-53, discloses:

“If the data to be accessed resides in the cache memory, then the port adapter accesses the data in the cache memory. If the data to be accessed does not reside in the cache memory, then the port adapter forwards a storage access request to the storage adapters 37, 38. One of

the storage adapters 37, 38 responds to the storage access request by performing a logical-to-physical translation to determine where the data to be accessed resides on the storage devices, and reads the data from the storage devices and writes the data to the cache memory, for access by the port adapter.” (emphasis added)

This section clearly indicates that the physical storage location depends on the data address and not on load balancing considerations. Further sections of Blumenau reinforce this section. For example, the examiner contends that column 12, lines 26-33 of Blumenau disclose a data storage request that includes a data identifier. However, this section of Blumenau actually discusses the contents of entries in a routing table that controls the switch. As stated there, each entry include a list pointer for the host, a storage port list for the host, a host name, a world-wide name for the host, and a source address (S_ID) for the host. All of this information refers to the host, not to data (see, for example Blumenau, column 12, lines 9-10).

Consequently, although data access requests can be routed along different paths between a host and the data storage subsystem, there is no teaching or suggestion in Blumenau that the data, to which the requests refer, could be stored in different storage devices based on load balancing considerations. Instead, data is statically assigned to specific storage devices. Therefore, the advantages achieved by the invention, such as automatic load balancing between storage devices and easy system scalability by adding storage devices without reconfiguration of the existing devices, cannot be achieved by the Blumenau system.

The examiner also indicates that applicants have argued that prior art did not disclose that data at an input node could be directed to different output nodes based on criteria other than the data address, such as load balancing considerations. This characterization of applicants' arguments is incorrect. The claims do not recite output nodes, but instead recite physical storage device locations in storage devices to which data is stored. The Blumenau sections to which the examiner refers (column 5, lines 58-67 and column 6, lines 1-5) disclose routing data to storage ports, which as discussed above, are not the same as physical device locations because data is not stored in a port, but passes through the port to be stored in a location in a storage device.

The Muller reference discloses a store and forward data packet switch in which data packets received at input ports are temporarily stored and then selectively forwarded to switch output ports based on information in a forwarding database. The bulk of the Muller disclosure is directed to hardware that allows a CPU to efficiently manage the information in the forwarding database as the switch operates. It would certainly be possible to combine the teachings of Blumenau and Muller. For example, the Muller switch could be used as the 32-port switch 40 disclosed in Blumenau. However, since neither of the cited references discloses a system in which data corresponding to the same data identifier or address may be stored in different physical storage device locations based on criteria, such as load balancing or availability and, consequently, the combination of these references cannot teach or suggest this type of operation.

The examiner contends “One [of] ordinary skill in the art at the time of the invention knows that ‘routing information based on the loading characteristic of the storage access request received at the switch inputs in order to balance loading of the storage access requests upon the outputs of the switch’ is prime example of load balancing, where the routing information is based on the loading characteristic of the meaning workload of the memory.” While applicants do not disagree with this contention, it is not what is being claimed. The claims are concerned with where the data is stored not how the data arrives there. For example, claim 1 (and claim 44 which has similar wording) recites:

“...an access interface module which receives from the client data storage requests, each including a data object to be stored and a data object identifier that identifies that data object and, in response to each service request and based on a workload and on relative demands placed on subsets of the plurality of storage devices instead of a physical location in the plurality of storage devices, dynamically selects a subset of the plurality of storage devices in which the data is stored so that data corresponding to the same data object identifier can be transferred to different physical storage device locations from request to request in order to dynamically distribute the workload across the plurality of storage devices.”

This claim wording does not read onto the Blumenau/Muller combination even if the entire Blumenau storage subsystem is considered to be a storage “device” because, under this interpretation, it would not be possible to select a subset of a plurality of storage devices as recited. It is also clear that, while the Blumenau/Muller system can alleviate overloading of the data storage subsystem network ports it cannot accomplish the objectives of the invention as set forth above.

The examiner claims that Blumenau discloses that data can be stored in different locations based on considerations, such as load balancing. However, the section of Blumenau to which the examiner refers (column 6, lines 46-53) discloses that the storage ports are adjusted based on load balancing concerns. As discussed above, data is not stored in storage ports. Instead, the data passes through the port to be stored at a location in a storage device. In Blumenau this storage device location is dictated by the address that is associated with the data and does not depend on considerations such as load balancing. Accordingly, Blumenau and Muller do not teach that data corresponding to the same data object identifier can be transferred to different physical storage device locations from request to request as recited in claims 1 and 44. Accordingly, claims 1 and 44 patentably distinguish over the cited reference combination.

Claims 3-11 and 46-54 are dependent, either directly or indirectly, on claims 1 and 44, respectively, and incorporate the limitations thereof. Therefore, they also distinguish over the cited reference combination in the same manner as claims 1 and 44. In addition, these claims recite further elements not disclosed or suggested by the cited reference combination.

Claims 3, 4, 5, 6, 46, 47, 48 and 49 recite details of the switch fabric that connects the access interface modules and the storage devices. These details include that the switch comprises separate control and data switch fabrics that are optimized for control and data signals, respectively; that the data is transmitted only through the data switch and that the data switch comprises a crosspoint switch and the control switch comprises an Ethernet switch. While it is clear that the Blumenau switch must switch both data requests and the data to which they refer, the switch itself is only shown as a single block (40) with no indication that two separate switches are used. The examiner

refers to Blumenau, column 2, lines 7-19 and 29-41. However, this section of Blumenau discusses how the switch computer routes data and requests between switch inputs and switch outputs; no internal switch details are provided. Therefore, the elements recited in claims 3 and 46 are not disclosed in Blumenau. Muller also discloses no switch analogous to the recitation of these claims.

Claims 4 and 47 recite that each switch half is optimized for transferring the information that it switches. The examiner cites Blumenau column 2, lines 29-41 as disclosing this limitation. However, as described above, this section of Blumenau discloses details of the computer that controls the switch and does not disclose details of the switch itself.

Claims 5 and 48 recite that the data and control information is separated by the access interface module. The examiner cites Blumenau column 6, lines 44-53, as disclosing this limitation. This section of Blumenau discloses a mechanism that adjusts the number of loop ports assigned to a host to achieve load balancing. However, the disclosed mechanism relates only to the data packets and does not disclose that control information is separated from the data as recited.

Claims 7 and 50 further recite a resource module connected to the plurality of storage devices that generates preallocation information that preallocates storage from the plurality of storage devices in order to evenly distribute a workload across the plurality of storage devices. As discussed above, Blumenau and Muller determine where data will be stored based on the data address; no preallocation is necessary or discussed. The examiner cites Blumenau, column 2, lines 5-21. This section of Blumenau discusses routing of data requests and, accordingly, no preallocation of any storage is discussed.

Claims 8 and 51 recite that the preallocation information is transferred to the access interface module through the switch. The examiner cites Muller, column 4, lines 44-57. This section of Muller discloses how Ethernet packets enter and leave a switch and how switches can be connected together to create a multi-layer switch. However, the preallocation information is recited in the parent claims 7 and 49 as information that that preallocates storage from the plurality of storage devices in order to evenly

distribute a workload across the plurality of resources. Muller does not mention preallocation of information at all.

Claims 9 and 52 recite that an access interface module selects a subset of the plurality of storage devices based on the preallocation information. As mentioned previously, the Blumenau and Muller references store data in storage devices at locations that are determined by addresses, not preallocation information. Therefore, these claims also distinguish over the cited reference combination.

Respectfully submitted

/paul e. kudirka/

Date: 2010-07-27

Paul E. Kudirka, Esq. Reg. No. 26,931
LAW OFFICES OF PAUL E. KUDIRKA
Customer Number 64967
Tel: (617) 357-0010 Fax: (617) 357-0035

VIII APPENDIX OF CLAIMS (37 C.F.R. §41.37(c)(1)(viii))

The text of the claims involved in the appeal is:

- 1 1. Apparatus for providing high-performance, scaleable data storage services to a
2 client from a plurality of storage devices, comprising:
3 an access interface module which receives from the client data storage
4 requests, each including a data object to be stored and a data object identifier that
5 identifies that data object and, in response to each storage request and based on a
6 workload and on relative demands placed on subsets of the plurality of storage
7 devices instead of a physical location in the plurality of storage devices, dynamically
8 selects a subset of the plurality of storage devices in which the data is stored so that
9 data corresponding to the same data object identifier can be transferred to different
10 physical storage device locations from request to request in order to dynamically
11 distribute the workload across the plurality of storage devices; and
12 a switch fabric for temporarily connecting the access interface module to
13 the selected subset of the plurality of storage devices so that the data can be
14 transferred to the selected subset of storage devices.
- 1 3. The apparatus of claim 1 wherein the switch fabric comprises a control switch
2 fabric for transferring control information and a separate data switch fabric for
3 transferring data.
- 1 4. The apparatus of claim 3 wherein the control switch fabric is optimized for
2 transferring control information and the data switch fabric is optimized for
3 transferring data.

1 5. The apparatus of claim 3 wherein the request for storage includes control
2 information and data and wherein the access interface module separates the
3 control information and the data and transfers the data to the selected subset of
4 storage devices over the data switch fabric.

1 6. The apparatus of claim 3 wherein the data switch fabric comprises a non-
2 blocking crossbar switch for data transfer and the control switch fabric comprises
3 an Ethernet switch for control information transfer.

1 7. The apparatus of claim 1 further comprising a resource module connected to the
2 plurality of storage devices for generating preallocation information that
3 preallocates storage from the plurality of storage devices in order to evenly
4 distribute a workload across the plurality of storage devices.

1 8. The apparatus of claim 7 wherein the switch fabric connects the access interface
2 module to the resource module so that the resource module can transfer the
3 preallocation information to the access interface module.

1 9. The apparatus of claim 8 wherein the access interface module selects a subset
2 of the plurality of storage devices based on the preallocation information.

1 10. The apparatus of claim 1 wherein the access interface module comprises a data
2 memory which temporarily stores information transferred between the access
3 interface module and the selected subset of the plurality of storage devices.

1 11. The apparatus of claim 1 further comprising a plurality of access interface
2 modules each access interface module receiving storage requests from a
3 plurality of clients.

1 44. A method for providing high-performance, scalable data storage services to a
2 client from a plurality of storage devices, the method comprising:

- 3 (a) providing an access interface module which receives from the client data
4 storage requests, each including a data object to be stored and a data
5 object identifier that identifies that data object;
- 6 (b) using the access interface module in response to each data storage
7 request and based on a workload and on relative demands placed on
8 subsets of the plurality of storage devices instead of a physical location in
9 the plurality of storage devices, dynamically selects a subset of the
10 plurality of storage devices in which the data is stored so that data
11 corresponding to the same data object identifier can be transferred to
12 different physical storage device locations from request to request in order
13 to dynamically distribute the workload across the plurality of storage
14 devices; and
- 15 (c) using a switch fabric to temporarily connect the access interface module to
16 the selected subset of the plurality of storage devices so that the data can
17 be transferred to the selected subset of storage devices.

1 46. The method of claim 44 wherein step (c) comprises:

- 2 (c1) using a control switch fabric for transferring control information; and

3 (c2) using a separate data switch fabric for transferring data.

1 47. The method of claim 46 wherein step (c1) comprises optimizing the control
2 switch fabric for transferring control information and step (c2) comprises
3 optimizing the data switch fabric for transferring data.

1 48. The method of claim 46 wherein the request for storage includes control
2 information and data and wherein step (b) comprises separating the control
3 information and the data and step (c) comprises transferring the data to the
4 selected subset of storage devices over the data switch fabric.

1 49. The method of claim 46 wherein step (c1) comprises using a non-blocking
2 crossbar switch for data transfer and step (c2) comprises using an Ethernet
3 switch for control information transfer.

1 50. The method of claim 44 further comprising:
2 (d) providing a resource module connected to the plurality of storage devices;
3 and
4 (e) using the resource module to generate preallocation information that
5 preallocates storage from the plurality of storage devices in order to
6 evenly distribute a workload across the plurality of storage devices.

1 51. The method of claim 50 wherein step (c) comprises connecting the access
2 interface module to the resource module so that the resource module can
3 transfer the preallocation information to the access interface module.

1 52. The method of claim 51 wherein step (b) comprises selecting a subset of the
2 plurality of storage devices based on the preallocation information.

1 53. The method of claim 44 wherein step (b) comprises temporarily storing
2 information transferred between the access interface module and the selected
3 subset of the plurality of storage devices.

1 54. The method of claim 44 wherein step (a) further comprises providing a plurality of
2 access interface modules each access interface module receiving storage
3 requests from a plurality of clients.

IX EVIDENCE APPENDIX (37 C.F.R. §41.37(c)(1)(ix))

None

X RELATED PROCEEDINGS APPENDIX (37 C.F.R. §41.37(c)(1)(x))

NONE